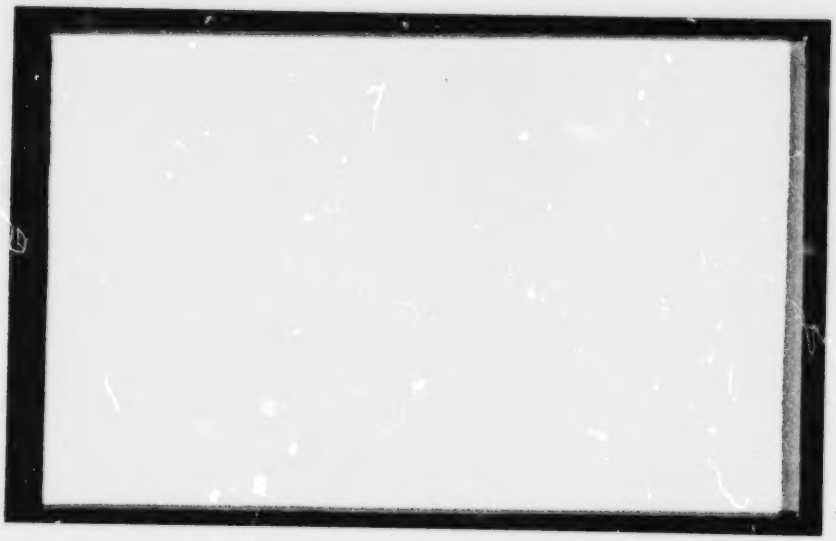
UNIVERSITÉ DE MONTRÉAL



CENTRE DE RECHERCHE SUR LES TRANSPORTS

MÉMOIRE D'ÉTUDIANT

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

MODELES DE DISTRIBUTION AVEC CONTRAINTES D'INEGALITE : CALIBRATION ET RESOLUTION

par
Robert Ducharme

Cette recherche a été soutenue par des subventions du Ministère de l'Education du Québec (Programme F.C.A.C.) et du Centre de recherche et de développement de Transports Canada.

Université de Montréal
Centre de recherche sur les transports - Publication # 170
Avril 1980

Dépôt légal - 2e trimestre 1980

TABLE DES MATIÈRES

SOMMAIRE iii

PRÉFACE iv

CHAPITRE I : INTRODUCTION 1

1. Historique 1
2. Maximisation de l'entropie et minimisation de l'information 2
3. La calibration et les théorèmes de Suzanne Evans . 6
4. Le modèle de distribution avec contraintes d'inégalité 8

CHAPITRE II: ALGORITHME DE BALANCEMENT GÉNÉRALISÉ 11

CHAPITRE III: CALIBRATION ET GÉNÉRALISATION DES THÉORÈMES DE S. EVANS 25

CHAPITRE IV: RÉSULTATS NUMÉRIQUES 38

1. Application du programme 4 38
2. Application du programme 5 46

CONCLUSION 52

REMERCIEMENTS 53

BIBLIOGRAPHIE 54

ANNEXE 1 : L'ALGORITHME DE BALANCEMENT 57

ANNEXE 2 : LA MÉTHODE D'APPROXIMATION LINÉAIRE DE HYMAN 58

ANNEXE 3 : L'ALGORITHME DE FRANK-WOLFE 59

SOMMAIRE

Les modèles de distribution basés sur la maximisation de l'entropie avec fonction de coût sont grandement utilisés dans divers problèmes d'interaction spatiale. L'une des raisons justifiant ce choix est l'efficacité de l'algorithme de balancement pour les résoudre. Une plus grande flexibilité peut être introduite dans ces modèles en permettant à certaines contraintes d'être sous signe d'inégalité. Le présent travail démontre qu'on peut généraliser l'algorithme de balancement pour calibrer et résoudre ces modèles ainsi modifiés.

PRÉFACE

Quatre types de programmes sont considérés dans ce travail, soit:

1- La maximisation de l'entropie avec une série de contraintes d'inégalité. La résolution de ce programme est donnée dans le chapitre II et sa calibration dans le chapitre III.

2- La minimisation de l'information avec une série de contraintes d'inégalité. Ce programme est résolu dans le chapitre II.

3- La maximisation de l'entropie avec une série de contraintes d'intervalle. La résolution et la calibration de ce programme sont considérées dans le chapitre IV.

4- La minimisation de l'information avec une série de contraintes d'intervalle. Ce programme est résolu dans le chapitre IV.

Voici des exemples de problèmes où les programmes 1, 2 et 3 s'appliquent.

Programme 1. L'exemple est tiré d'un article de Florian et Los [8] et se situe dans un contexte de "park'n ride". Le problème considéré est la prévision de l'impact de certaines politiques sur le choix des gares par les usagers, qui stationnent leur automobile à la gare et se rendent au centre-ville en train. Florian et Los démontrent que la modélisation basée sur le programme 1 est celle qui réflète le mieux l'interaction entre l'offre et la demande. Soient:

O_i = Nombre d'usagers qui partent de l'origine i

D_j = Nombre de personnes que peut accueillir le parc de stationnement de la gare j (en tenant compte de l'occupation moyenne des automobiles stationnées)

c_{ij} = coût généralisé impliqué par le choix de la gare j pour un usager de la zone i

T_{ij} = Nombre d'usagers de la zone i qui choisissent la gare j .

La distribution la plus probable de cette dernière variable est celle qui résout le programme suivant:

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_{j} T_{ij} = O_i \quad \forall i$$

$$\sum_{i} T_{ij} \leq D_j \quad \forall j$$

$$T_{ij} \geq 0 \quad \forall (i,j) .$$

La demande est directement reflétée dans la matrice de coûts $[c_{ij}]$ tandis que les contraintes d'offre sont explicites.

Cet exemple est discuté dans le chapitre V.

Programme 2. Supposons qu'on veuille prévoir le comportement du marché du logement dans un contexte qui n'est pas celui de la concurrence pure (contrôle des loyers). Supposons également un stock de logements excédant la demande.

La prévision pourrait se faire de la façon suivante:

1- Prévoir la demande "ex ante" (demande Walrasienne) de logements par les ménages sans tenir compte des contraintes d'offre.

2- Trouver la distribution "ex post" compatible avec les contraintes d'offre la plus près possible (au sens de la théorie de l'information) de la distribution "ex ante". Cette seconde étape est réalisée au moyen du programme 2.

Soient:

O_i = Nombre de ménages de type i

D_j = Nombre de logements dans la zone j

A_{ij} = Prédiction "ex ante" de la demande de logements dans la zone j par les ménages de type i

P_{ij} = Prédiction "ex post" de la demande de logements dans la zone j par les ménages de type i .

Naturellement,

$$\sum_j A_{ij} = O_i \quad \forall i \ .$$

La distribution "ex post" P_{ij} sera la solution du programme suivant:

Minimiser

$$\sum_{i,j} P_{ij} \ln \left(\frac{P_{ij}}{A_{ij}}\right)$$

sujet à:

$$\sum_{j} P_{ij} = O_i \quad \forall i$$

$$\sum_{i} P_{ij} \leq D_j \quad \forall j$$

$$P_{ij} \geq 0 \quad \forall (i,j) .$$

(N.B. Si aucune des contraintes d'offre n'est contraignante, la distribution "ex post" sera égale à la distribution "ex ante".)

Programme 3. Un exemple typique de l'utilisation de ce programme est celui où l'on dispose de données inconsistantes. Situons-nous dans un contexte où l'on veut prévoir le commerce d'une certaine marchandise entre les différentes régions d'un pays à partir des données suivantes:

O_i = Exportations en tonnes de la région i aux autres régions

D_j = Importations en tonnes de la région j à partir des autres régions

c_{ij} = Coût d'une expédition d'une tonne de la région i à la région j.

Normalement, on devrait avoir $\sum_i O_i = \sum_j D_j$.

Supposons que les O_i sont des données fiables alors que les D_j

sont incertains et que $\sum_i O_i \neq \sum_j D_j$.

On peut alors définir un intervalle de confiance $[\underline{D}_j, \overline{D}_j]$ pour chaque D_j de telle sorte que:

$$\sum_j \underline{D}_j < \sum_i O_i < \sum_j \overline{D}_j \ .$$

En résolvant le programme suivant:

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_j T_{ij} = O_i \quad \forall i$$

$$\underline{D}_j \leq \sum_i T_{ij} \leq \overline{D}_j \quad \forall_j$$

$$T_{ij} \geq 0 \quad \forall (i,j)$$

on estime non seulement le commerce entre les régions mais encore l'importation de chacune des régions.

CHAPITRE 1

INTRODUCTION

1. Historique

Le modèle de distribution basé sur la maximisation de l'entropie fut introduit par Wilson en 1970 [23]. Cette modélisation peut être justifiée de deux façons:

a) Par une théorie objective des probabilités: Wilson a démontré que la distribution qui a le plus de chances de se réaliser est celle qui maximise l'entropie (à approximation de Stirling près).

b) Par une théorie subjective des probabilités, fondée sur le degré de croyance dans un événement. La seule mesure d'incertitude d'une distribution, à facteur multiplicatif près, est son entropie. La distribution qui maximise l'entropie est donc celle qui donne le moins d'information. Cette justification, également donnée par Wilson, est tirée de l'interprétation de Jaynes [12] de la théorie de l'information, inventée par Shannon [18].

En 1977, Snickars et Weibull [20] ont ajouté une extension à l'ap-

proche classique de la maximisation de l'entropie: il s'agit du principe de minimisation de la transmission d'information entre deux distributions de probabilité. Ce principe permet l'utilisation d'information à priori sur la distribution lorsqu'on en dispose. S'il n'y a aucune information à priori, c'est-à-dire si toutes les interactions sont équiprobables à priori, on retrouve la maximisation de l'entropie.

2. *Maximisation de l'entropie et minimisation de l information*

Les modèles de distribution visent à estimer les échanges entre les différents couples origine-destination dans un système d'interaction spatiale. Cette estimation se fait à partir de l'information dont on dispose et qui ne porte en général que sur des états agrégés du système. Par exemple, dans le cadre d'une économie nationale, on connaît les exportations et les importations des différentes régions mais on ignore le commerce entre ces régions. Ou encore, dans le contexte d'une ville close, on connaît le nombre de travailleurs habitant les différentes zones résidentielles et le nombre de travailleurs qui ont un emploi dans les différentes zones commerciales, mais on ignore les interactions entre zones résidentielles et commerciales.

Il est raisonnable de choisir comme estimation la matrice d'interactions qui a le plus de chances de se réaliser parmi celles qui sont compatibles avec l'information partielle dont on dispose. Soient:

O_i = Nombre de travailleurs habitant la zone i

D_j = Nombre d'individus travaillant dans la zone j

T = Nombre total de travailleurs $(T = \sum_i O_i = \sum_j D_j)$

x_k = Spécification de la zone de résidence et de travail de l'individu k

T_{ij} = Nombre de travailleurs qui résident dans la zone i et travaillent dans la zone j .

Les O_i , D_j et T sont l'information partielle dont on dispose; les T_{ij} sont à estimer.

$\underline{x} = (x_1, x_2, \ldots, x_T)$ est un microétat. Il s'agit d'une spécification complète et détaillée du système.

$\underline{T} = \begin{pmatrix} T_{11} & \cdots & T_{1J} \\ \vdots & & \vdots \\ T_{I1} & \cdots & T_{IJ} \end{pmatrix}$ est un macroétat. Nous voulons estimer le macroétat le plus probable.

Un macroétat satisfaisant:

$$\sum_{j=1}^{J} T_{ij} = O_i \quad \forall\, i \tag{1}$$

$$\sum_{i=1}^{I} T_{ij} = D_j \quad \forall\, j \tag{2}$$

$$T_{ij} \in \mathbb{N} \quad \text{où} \quad \mathbb{N} = \{0,1,2,\ldots\} \quad \forall\, (i,j) \tag{3}$$

est réalisable. Tous les microétats contenus dans un macroétat réalisable sont également réalisables. L'hypothèse fondamentale suivante permet d'établir un critère pour déterminer le macroétat le plus probable:

Hypothèse: Tout microétat non-réalisable a une probabilité nulle et tous les microétats réalisables sont équiprobables.

Il s'ensuit que tout macroétat non-réalisable a une probabilité nulle et que la probabilité d'un macroétat réalisable est proportionnelle au nombre de microétats qu'il contient. Le macroétat le plus probable est donc celui qui contient le plus de microétats parmi ceux qui satisfont les équations (1), (2) et (3). Il est la solution du programme:

Maximiser

$$\frac{T!}{\prod_{i,j} T_{ij}\,!}$$

sujet à (1), (2), (3).

En maximisant le logarithme naturel et en utilisant l'approximation de Stirling ($K! \approx K^K e^{-K}$) on obtient le programme mathématique suivant:

Programme 1.

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij}$$

sujet à (1), (2) et

$$T_{ij} \geq 0 \quad \forall(i,j) \ . \tag{4}$$

On maximise donc l'entropie de la distribution sous les contraintes (1), (2) et (4). La solution de ce programme est:

$$T_{ij} = \frac{O_i \ D_j}{T} \ \forall(i,j) \ .$$

Si on dispose en plus d'une information sur le coût d'un échange entre i et j, c_{ij}, le programme à résoudre est le suivant:

Programme 2.

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} \ T_{ij}$$

sujet à (1), (2), (4).

Le paramètre θ est un arbitre entre les deux objectifs que sont la maximisation de l'entropie et la minimisation du coût global des échanges. La solution de ce programme est de la forme:

$$T_{ij} = Q_i \ P_j \ \exp(-\theta \ c_{ij}) \ \forall(i,j)$$

i.e. la matrice $[T_{ij}]$ est biproportionnelle à la matrice $[\exp(-\theta c_{ij})]$. (Les Q_i et P_j sont uniques à facteur multiplicatif près).

Considérons par ailleurs le cas où l'on dispose d'une information à priori B_{ij} quant à l'interaction entre les zones i et j. Il

s'agit alors de trouver la matrice $[T_{ij}]$ qui minimise le gain d'information par rapport à la matrice $[B_{ij}]$ tout en satisfaisant les contraintes (1), (2) et (4). Le programme à résoudre est le suivant:

Programme 3.

Minimiser

$$\sum_{i,j} T_{ij} \ln\left(\frac{T_{ij}}{B_{ij}}\right)$$

sujet à (1), (2), (4).

La solution est de la forme:

$$T_{ij} = Q_i \, P_j \, B_{ij} \quad \forall(i,j)$$

i.e. la matrice $[T_{ij}]$ est biproportionnelle à la matrice $[B_{ij}]$.

On remarque que:

1- Si on ne dispose d'aucune information à priori, on doit supposer toutes les interactions équiprobables à priori. On a alors $B_{ij} = 1 \quad \forall(i,j)$ et on retrouve le programme 1.

2- Si $B_{ij} = \exp(-\theta \, c_{ij}) \quad \forall(i,j)$, on retrouve le programme 2.

3. *La calibration et les théorèmes de Suzanne Evans.*

L'estimation du paramètre θ dans le programme 2 constitue la calibration du modèle. Cette estimation se fait sur une période où l'on dispose de la matrice d'interactions observées, et on suppose θ constant durant la période de prévision.

Si on permet au paramètre θ de varier dans le programme 2, la matrice d'interactions optima'e qui solutionne ce programme sera elle-même fonction de θ . Andrew Evans [5] a démontré que l'estimateur de vraisemblance maximale de θ est celui pour lequel le coût global des interactions de la matrice qui solutionne le programme 2 est égal au coût global des interactions observées durant la même période. Il a été de plus démontré ([5], [6]) que cet estimateur doit être unique, en vertu de la décroissance stricte de la fonction qui relie le coût global des interactions de la matrice qui solutionne le programme 2 au paramètre θ , à l'exception d'un cas trivial où cette matrice est indépendante de θ .

Suzanne Evans [6] a de plus démontré le théorème suivant, qui assure l'existence de l'estimateur du paramètre θ :

Théorème. Le coût global des interactions de la matrice qui solutionne le programme 2, lorsque $\theta \to \infty$, tend vers la valeur à l'optimum de la fonction-objectif du programme suivant:

Minimiser

$$\sum_{i,j} c_{ij} T_{ij}$$

sujet à (1), (2), (4).

Symétriquement, ce théorème demeure exact si l'on remplace ∞ par $-\infty$ et minimiser par maximiser.

Ces caractéristiques font que le programme 2 est assez facile à calibrer. Plusieurs méthodes de calcul de l'estimateur de θ sont possibles, toutes basées sur la résolution successive du programme 2 à partir de quelques valeurs de θ. La méthode d'approximation linéaire de Hyman est une de ces méthodes, et est donnée en annexe 3.

4. *Le modèle de distribution avec contraintes d'inégalité*.

Une plus grande flexibilité peut être introduite dans les modèles précédents en permettant à certaines contraintes d'être sous signe d'inégalité. Dans le cas de la ville, par exemple, il peut y avoir des gens de l'extérieur de la ville qui travaillent dans les zones commerciales. On a alors $T = \sum_i O_i < \sum_j D_j$. Le programme à résoudre est le suivant:

Programme 4.

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_j T_{ij} = O_i \quad \forall i \tag{5}$$

$$\sum_i T_{ij} \leq D_j \quad \forall j \tag{6}$$

$$T_{ij} \geq 0 \quad \forall (i,j) \tag{7}$$

(Le problème serait exactement le même si on avait capacité excédentaire des origines plutôt que des destinations.)

Un autre cas est celui où on dispose de prévisions ou de données incertaines quant aux D_j ; on peut alors définir un intervalle pour les D_j et le programme à résoudre est:

Programme 5.

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_{j} T_{ij} = O_i \quad \forall i \tag{8}$$

$$\sum_{i} T_{ij} \leq \overline{D}_j \quad \forall j \tag{9}$$

$$\sum_{i} T_{ij} \geq \underline{D}_j \quad \forall j \tag{10}$$

$$T_{ij} \geq 0 \quad \forall (i,j) \; . \tag{11}$$

Deux hypothèses non-restrictives sont faites sur les programmes 4 et 5:

1- $\exists u$ et v tels que $c_{ij} = u_i + v_j \quad \forall (i,j)$

2- Les déplacements observés lors de la période de calibration ne solutionnent ni le programme de maximisation ni le programme de minimisation du coût global des interactions.

La section 2 généralise l'algorithme de balancement pour résoudre le programme 4. La section 3 généralise les théorèmes de Suzanne Evans au programme 4 et indique comment calibrer ce modèle.

La section 4 généralise les résultats des sections 2 et 3 au programme 5.

La section 5 présente des tests numériques effectués sur les programmes 4 et 5 avec une insistance particulière sur la rapidité des temps de résolution.

La section 6 conclut.

CHAPITRE II

ALGORITHME DE BALANCEMENT GÉNÉRALISÉ

Le but de cette section est la résolution du programme 4:

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_{j} T_{ij} = O_i \quad \forall i \tag{5}$$

$$\sum_{i} T_{ij} \leq D_j \quad \forall j \tag{6}$$

$$T_{ij} \geq 0 \quad \forall (i,j) \; . \tag{7}$$

Enonçons d'abord les notations et conventions utilisées au cours des chapitres II et III: $\psi(\underset{\sim}{T},\theta) = F(\underset{\sim}{T}) + \theta\, C(\underset{\sim}{T})$ est la fonction-objectif, où

$$F(\underset{\sim}{T}) = \sum_{i,j} T_{ij} \ln T_{ij}$$

$$C(\underset{\sim}{T}) = \sum_{i,j} c_{ij} T_{ij}$$

$\underset{\sim}{T}(\theta)$ est la solution du programme 4, étant donné θ .

$\hat{\psi}(\theta) = \psi(\underset{\sim}{T}(\theta),\theta) = \hat{F}(\theta) + \theta\,\hat{C}(\theta)$ est la valeur optimale de la fonction-objectif étant donné θ , où

$$\hat{F}(\theta) = F(\underset{\sim}{T}(\theta)) = \sum_{i,j} T_{ij}(\theta) \ln T_{ij}(\theta)$$

$$\hat{C}(\theta) = C(\underset{\sim}{T}(\theta)) = \sum_{i,j} c_{ij}\, T_{ij}(\theta)$$

$$g^t(\underset{\sim}{T}) = (g_1(\underset{\sim}{T}),\ldots,g_I(\underset{\sim}{T})) \quad \text{où} \quad g_i(\underset{\sim}{T}) = \sum_j T_{ij} - O_i$$

$$h^t(\underset{\sim}{T}) = (h_1(\underset{\sim}{T}),\ldots,h_J(\underset{\sim}{T})) \quad \text{où} \quad h_j(\underset{\sim}{T}) = \sum_i T_{ij} - D_j$$

R^K = espace euclidien de dimension K

∇ = gradient .

Dans le cas d'un vecteur $a^t = (a_1,\ldots,a_K)$ nous utiliserons la convention suivante:

$$a > 0 \rightarrow a_i > 0 \quad \forall i$$

$$a \geq 0 \rightarrow a_i \geq 0 \quad \forall i$$

$$a \geq 0 \rightarrow a_i \geq 0 \quad \forall i \quad \text{et} \quad a_i > 0 \quad \text{pour au moins un} \quad i .$$

Nous supposerons θ connu dans la section 2 et nous écrirons $\psi(\underset{\sim}{T})$ plutôt que $\psi(\underset{\sim}{T},\theta)$. Le programme 4 s'écrit maintenant:

Minimiser

$$\psi(\underline{T})$$

sujet à:

$$g(\underline{T}) = 0 \tag{12}$$

$$h(\underline{T}) \leq 0 \tag{13}$$

$$\underline{T} \geq 0 \ . \tag{14}$$

Il s'agit d'un programme de minimisation d'une fonction non-linéaire sous des contraintes linéaires. La contrainte (14) définit le domaine de la fonction ψ, qui est un sous-ensemble de $\mathbf{R}^{IJ}$. On inclut $T_{ij} = 0$ dans le domaine en fixant $T_{ij} \ln T_{ij} = 0$ lorsque $T_{ij} = 0$, ce qui est le résultat à la limite. Les contraintes (12) et (13) définissent le sous-ensemble réalisable du domaine.

L'ensemble réalisable Ω induit par (12), (13), (14) est un ensemble convexe et ψ est une fonction strictement convexe sur Ω. Le fait que Ω soit un ensemble fermé et borné assure l'existence d'une solution du programme 4.

Théorème 1. La solution optimale d'un programme de minimisation d'une fonction strictement convexe sur un ensemble convexe, si elle existe, est unique.

Preuve. Voir Mangasarian [17]. □

Théorème 2. Le programme 4 est équivalent au problème de Kuhn-Tucker suivant: trouver

$$\bar{\underset{\sim}{T}} \in R^{IJ} , \bar{\underset{\sim}{T}} \geq 0 , \bar{\alpha} \in \mathbb{R}^I , \bar{\beta} \in \mathbb{R}^J , \bar{\beta} \geq 0 ,$$

s'ils existent, tels que:

$$L(\bar{\underset{\sim}{T}},\alpha,\beta) \leq L(\bar{\underset{\sim}{T}},\bar{\alpha},\bar{\beta}) \leq L(\underset{\sim}{T},\bar{\alpha},\bar{\beta}) \; \forall \underset{\sim}{T} \in \mathbb{R}^{IJ} , T \geq 0 , \alpha \in \mathbb{R}^I , \beta \in \mathbb{R}^J , \beta \geq 0$$

où

$$\begin{aligned} L(\underset{\sim}{T},\alpha,\beta) &= \psi(\underset{\sim}{T}) + \alpha^t g(\underset{\sim}{T}) + \beta^t h(\underset{\sim}{T}) \\ &= \sum_{i,j} T_{ij}(\ln T_{ij} + \theta c_{ij}) + \sum_i \alpha_i (\sum_j T_{ij} - O_i) \\ &\quad + \sum_j \beta_j (\sum_i T_{ij} - D_j) . \end{aligned}$$

On a de plus les relations de complémentarité suivantes:

$$\bar{\beta}_j > 0 \rightarrow h_j(\bar{\underset{\sim}{T}}) = 0 \leftrightarrow \sum_j \bar{T}_{ij} = D_j$$

$$\sum_i \bar{T}_{ij} < D_j \leftrightarrow h_j(\underset{\sim}{T}) < 0 \rightarrow \bar{\beta}_j = 0 .$$

Preuve. Voir Mangasarian [17] . □

La fonction L est appelée fonction de Lagrange et les variables α_i et β_j multiplicateurs de Lagrange ou variables duales associés aux contraintes $g_i(\underset{\sim}{T}) = 0$ et $h_j(\underset{\sim}{T}) \leq 0$ respectivement. La matrice

$\underset{\sim}{T}$ contient les variables primales. Ce problème est un problème de "point-selle" en ce sens que la fonction L est maximisée relativement aux variables duales et minimisée relativement aux variables primales. La solution $(\overline{\underset{\sim}{T}},\overline{\alpha},\overline{\beta})$ est unique. (Dans le cas du modèle de distribution à contraintes d'égalité, cela n'est pas vrai: en effet, le fait que $\sum_i O_i = \sum_j D_j$ permet d'additionner $\lambda \in \mathbb{R}$ à tous les α_i et de soustraire λ de tous les β_j sans changer la valeur de L. Cela est dû au fait qu'il y a alors dépendance linéaire entre les contraintes).

Le théorème dit que $\overline{\underset{\sim}{T}}$ est la solution optimale du programme 4 si et seulement si $(\overline{\underset{\sim}{T}},\overline{\alpha},\overline{\beta})$ est la solution optimale du problème de Kuhn-Tucker. Ce théorème est valable même si la fonction L n'est pas différentiable, sous certaines conditions. Lorsque la fonction L est différentiable, ce qui est le cas ici, le problème de Kuhn-Tucker peut être énoncé de la façon suivante:

Trouver $\overline{\underset{\sim}{T}} \in R^{IJ}$, $\overline{\underset{\sim}{T}} \geq 0$, $\overline{\alpha} \in \mathbb{R}^I$, $\overline{\beta} \in \mathbb{R}^J$, s'ils existent, tels que:

$$\nabla_{\underset{\sim}{T}} L(\overline{\underset{\sim}{T}},\overline{\alpha},\overline{\beta}) = 0$$

$$\nabla_{\alpha} L(\overline{\underset{\sim}{T}},\overline{\alpha},\overline{\beta}) = 0$$

$$\nabla_{\beta} L(\overline{\underset{\sim}{T}},\overline{\alpha},\overline{\beta}) \leq 0$$

$$\overline{\beta}\, \nabla_{\beta} L(\overline{\underset{\sim}{T}},\overline{\alpha},\overline{\beta}) = 0$$

$$\overline{\beta} \geq 0$$

Dans notre problème cela signifie:

$$\bar{T}_{ij} = \exp(-1-\bar{\alpha}_i)\ \exp(-\bar{\beta}_j)\ \exp(-\theta\ c_{ij}) \quad \forall(i,j) \tag{15}$$

$$\sum_j \exp(-1-\bar{\alpha}_i)\ \exp(-\bar{\beta}_j)\ \exp(-\theta\ c_{ij}) = O_i \quad \forall i \tag{16}$$

$$\sum_i \exp(-1-\bar{\alpha}_i)\ \exp(-\bar{\beta}_j)\ \exp(-\theta\ c_{ij}) \leq D_j \quad \forall j \tag{17}$$

$$\bar{\beta}_j \sum_i \exp(-1-\bar{\alpha}_i)\ \exp(-\bar{\beta}_j)\ \exp(-\theta\ c_{ij}) = \bar{\beta}_j\ D_j \quad \forall_j \tag{18}$$

$$\bar{\beta}_j \geq 0 \quad \forall j\ . \tag{19}$$

Les équation (17) à (19) sont les relations de complémentarité. L'ensemble des équations (15) à (19) constitue les conditions de Kuhn-Tucker.

Nous allons maintenant résoudre le problème de Kuhn-Tucker. Il faut solutionner:

$$\left.\begin{array}{l}\text{maximiser} \\ \alpha \text{ quelconque} \\ \beta \geq 0\end{array}\right\} \left\{\begin{array}{ll}\text{minimiser} & L(\underset{\sim}{T},\alpha,\beta) \\ \underset{\sim}{T} \geq 0 & \end{array}\right\}$$

Fixons arbitrairement α et β. L est alors une fonction strictement convexe que nous devons minimiser sans contraintes ($\underset{\sim}{T} \geq 0$ est une contrainte redondante). L'annulation des dérivées partielles est une condition nécessaire et suffisante pour obtenir le minimum unique de L :

$$\nabla_{\underset{\sim}{T}} L(\underset{\sim}{T},\alpha,\beta) = 0 \iff T_{ij} = \exp(-1-\alpha_i)\exp(-\beta_j)\exp(-\theta c_{ij}) \quad \forall(i,j)$$

$$\iff \underset{\sim}{T} = \underset{\sim}{T}(\alpha,\beta) \ . \tag{20}$$

En substituant $\underset{\sim}{T}$ dans L par l'équation (20), nous obtenons le programme suivant:

Programme 4^D.

Maximiser

$$L(\underset{\sim}{T}(\alpha,\beta),\alpha,\beta) = -\sum_i \alpha_i O_i - \sum_j \beta_j D_j - \sum_{i,j} \exp(-1-\alpha_i)\exp(-\beta_j)\exp(-\theta c_{ij})$$

sujet à $\beta_j \geq 0 \quad \forall j$. (21)

Il s'agit d'un programme de maximisation d'une fonction strictement concave sur un ensemble convexe.

Définition 3. Le programme dual associé au programme 4 (programme primal) est le suivant:

Maximiser

$$\underset{\sim}{L}(T,\alpha,\beta)$$

sujet à $(\underset{\sim}{T},\alpha,\beta) \in Y$ où

$$Y = \{(\underset{\sim}{T},\alpha,\beta)/\underset{\sim}{T}\in R^{IJ}, \underset{\sim}{T}\geq 0, \alpha\in R^I, \beta\in R^J, \nabla_{\underset{\sim}{T}} L(\underset{\sim}{T},\alpha,\beta)=0, \beta\geq 0\} \ .$$

Voir Mangasarian [17]. □

Soit:

$$Z = \{(\alpha,\beta)/\alpha \in R^I \ , \ \beta \in R^J \ , \ \beta \geq 0\} \ ;$$

alors:

$$(\underset{\sim}{T},\alpha,\beta) \in Y \iff (\alpha,\beta) \in Z \quad \text{et} \quad \underset{\sim}{T} = \underset{\sim}{T}(\alpha,\beta) \ .$$

Le programme 4^D est donc le programme dual associé au programme 4.

Théorème 4. La solution optimale d'un programme de maximisation d'une fonction strictement concave sur un ensemble convexe, si elle existe, est unique.

Preuve. Voir Mangasarian [17]. □

Théorème 5. Théorème de dualité faible. Soient

$$\underset{\sim}{T}^1 \in \Omega \quad \text{et} \quad (\alpha^2,\beta^2) \in Z \ .$$

Alors:

$$\psi(\underset{\sim}{T}^1) \geq L(\underset{\sim}{T}(\alpha^2,\beta^2),\alpha^2,\beta^2) \ .$$

Preuve. Voir Mangasarian [16]. □

Théorème 6. Théorème de dualité forte. Soit $\overline{\underset{\sim}{T}}$ la solution optimale du programme 4. Alors il existe $(\overline{\alpha},\overline{\beta})$ tels que:

$$\psi(\bar{\underset{\sim}{T}}) = L(\underset{\sim}{T}(\bar{\alpha},\bar{\beta}),\bar{\alpha},\bar{\beta})$$

$(\bar{\alpha},\bar{\beta})$ est la solution optimale du programme 4^D et $\bar{\underset{\sim}{T}} = \underset{\sim}{T}(\bar{\alpha},\bar{\beta})$.

Preuve. Voir Magasarian [16]. □

Soit $(\alpha,\beta) \in Z$. Nous avons:

$$\psi(\underset{\sim}{T}(\alpha,\beta))) = -\sum_i \alpha_i \sum_j \exp(-1-\alpha_i)\exp(-\beta_j)\exp(-\theta\, c_{ij})$$

$$-\sum_j \beta_j \sum_i \exp(-1-\alpha_i)\exp(-\beta_j)\exp(-\theta\, c_{ij})$$

$$-\sum_{i,j} \exp(-1-\alpha_i)\exp(-\beta_j)\exp(-\theta\, c_{ij}) \qquad (22)$$

$$L(\underset{\sim}{T}(\alpha,\beta),\alpha,\beta) = -\sum_i \alpha_i O_i - \sum_j \beta_j D_j$$

$$-\sum_{i,j} \exp(-1-\alpha_i)\exp(-\beta_j)\exp(-\theta\, c_{ij}) \; . \qquad (23)$$

Si $(\alpha,\beta) = (\bar{\alpha},\bar{\beta})$, ces deux expressions sont égales d'après le théorème 6; cela peut également se vérifier à partir des conditions de Kuhn-Tucker.

Nous allons maintenant développer un algorithme pour résoudre le programme 4^D (cet algorithme a déja été développé par Jefferson et Scott [13], mais par le biais de la programmation géométrique).

Remarquons d'abord que le théorème 6 assure l'existence d'une solution au programme 4^D (car il existe une (et une seule) solution au programme 4, comme nous l'avons vu). Le théorème 4 nous assure de l'unicité de cette solution.

Le mécanisme de l'algorithme sera le suivant: nous allons d'abord assigner au vecteur β une valeur se situant dans l'ensemble réalisable $Z(\beta^1 = 0)$; nous allons ensuite chercher α^1 tel que

$$L(\underset{\sim}{T}(\alpha^1,\beta^1),\alpha^1,\beta^1) = \text{Maximum}\{L(T(\alpha,\beta^1),\alpha,\beta^1)/(\alpha,\beta^1) \in Z\} .$$

Nous verrons plus loin que α^1 existe et est unique. α^1 étant fixé, nous allons chercher β^2 tel que

$$L(\underset{\sim}{T}(\alpha^1,\beta^2),\alpha^1,\beta^2) = \text{Maximum}\{L(T(\alpha^1,\beta),\alpha^1,\beta)/(\alpha^1,\beta) \in Z\} .$$

Nous verrons ici aussi que β^2 existe et est unique. Comme $(\alpha^1,\beta^1) \in Z$, nous aurons:

$$L(\underset{\sim}{T}(\alpha^1,\beta^2),\alpha^1,\beta^2) \geq L(\underset{\sim}{T}(\alpha^1,\beta^1),\alpha^1,\beta^1)$$

avec égalité si et seulement si $\beta^2 = \beta^1$.

β^2 étant maintenant fixé, nous chercherons α^2 tel que

$$L(\underset{\sim}{T}(\alpha^2,\beta^2),\alpha^2,\beta^2) = \text{Maximum}\{L(\underset{\sim}{T}(\alpha,\beta^2),\alpha,\beta^2)/(\alpha,\beta^2) \in Z\}$$

et ainsi de suite... Nous obtiendrons alors la chaîne suivante:

$$L(\underset{\sim}{T}(\alpha^1,\beta^1),\alpha^1,\beta^1) \leq L(\underset{\sim}{T}(\alpha^1,\beta^2),\alpha^1,\beta^2) \leq L(\underset{\sim}{T}(\alpha^2,\beta^2),\alpha^2,\beta^2) \leq \ldots$$

telle que

$$\{(\alpha^n,\beta^n)\ ,\ (\alpha^n,\beta^{n+1})\} \in Z \quad \forall\, n \in \mathbb{N} - \{0\}\ .$$

Cette suite converge puisque L est bornée supérieurement sur Z. Nous allons d'abord expliciter le calcul de cette suite puis montrer qu'elle converge vers la solution du programme 4^D.

1- β^1 étant fixé, $L(\underset{\sim}{T}(\alpha,\beta^1),\alpha,\beta^1)$ est une fonction strictement concave, bornée supérieurement et sans contraintes de α. Son maximum, s'il existe, est atteint en:

$$\nabla_\alpha\, L(\underset{\sim}{T}(\alpha,\beta^1),\alpha,\beta^1) = 0 \iff \exp(-1-\alpha_i^1) = \frac{O_i}{\sum_j \exp(-\beta_j^1)\ \exp(-\theta\ c_{ij})} \quad \forall i$$

$\therefore$ α^1 existe et est unique.

2- α^1 étant maintenant fixé, nous allons calculer β^2. Nous allons maximiser L successivement par rapport à $\beta_1,\beta_2,\ldots$ en supposant les autres β_j fixés à la dernière valeur qui leur a été assignée. Nous verrons que ce procédé est équivalent à la maximisation de L par rapport au vecteur β.

Soit β_1 variable, $\beta_j = \beta_j^1$ $\forall j \neq 1$. L est alors une fonction de β_1, strictement concave et bornée supérieurement, à maximiser sur l'ensemble $\beta_1 \geq 0$. Son maximum, s'il existe, sur $\beta_1 \in \mathbb{R}$, est atteint en:

$$\exp(-\beta_1) = \frac{D_1}{\sum_i \exp(-1-\alpha_i^1)\exp(-\theta\, c_{ij})}$$

$\therefore$ cette valeur de β_1 existe et est unique. Maintenant, si la valeur à droite de l'égalité est inférieure ou égale à 1, nous avons $\beta_1 \geq 0$. Si cette valeur est supérieure à 1, le maximum de L est atteint en $\beta_1 < 0$. L étant strictement concave, le maximum de L sur l'ensemble $\beta_1 \geq 0$ est atteint en $\beta_1 = 0$. Le maximum de L sur l'ensemble $\beta_1 \geq 0$ est donc donné par l'expression:

$$\exp(-\beta_1^2) = \text{Minimum}\left(1, \frac{D_1}{\sum_i \exp(-1-\alpha_i^1)\exp(-\theta\, c_{i1})}\right) .$$

En recommençant cette procédure pour $j = 2,3,\ldots,J$ nous avons:

$$\exp(-\beta_j^2) = \text{Minimum}\left(1, \frac{D_j}{\sum_i \exp(-1-\alpha_i^1)\exp(-\theta\, c_{ij})}\right) \quad \forall j$$

Nous voyons donc que chacun des β_j n'est fonction que de α^1 . Par conséquent, la maximisation séquentielle de L par rapport à $\beta_1,\beta_2,\ldots$ est équivalente à la maximisation de L par rapport à β .

3- β^2 étant maintenant fixé, nous répétons l'opération et obtenons:

$$\exp(-1-\alpha_i^2) = \frac{O_i}{\sum_j \exp(-\beta_j^2)\exp(-\theta\, c_{ij})} \quad \forall i$$

etc...

Nous remarquons que ce processus satisfait tour à tour certaines conditions de Kuhn-Tucker: après les étapes d'ordre impair (1,3,...) l'équation (16) est satisfaite; après les étapes d'ordre pair (2,4,...) les équations (17), (18) et (19) sont satisfaites.

Il s'ensuit que le point vers lequel la suite $\{(\alpha^n,\beta^n),(\alpha^n,\beta^{n+1})\}$ converge satisfait simultanément toutes les conditions de Kuhn-Tucker; il s'agit de $(\bar{\alpha},\bar{\beta})$. Ceci termine la preuve de la convergence de l'algorithme.

Lorsque le programme 4^D a été solutionné, on a aussi obtenu la solution du programme 4 qui peut être déduite de la relation (20).

Nous sommes maintenant en mesure d'énoncer d'une façon précise les étapes qui nous mèneront à la résolution simultanée des programmes 4^D et 4 (nous posons $Q_i = \exp(-1-\alpha_i)$ et $P_j = \exp(-\beta_j)$ et nous établissons un critère d'arrêt approprié):

1- $n = 1$

2- $P_j^n = 1 \quad \forall j$

3- $$Q_i^n = \frac{O_i}{\sum_j P_j^n \exp(-\theta c_{ij})} \quad \forall i$$

4- Si $P_j^n \sum_i Q_i^n \exp(-\theta\, c_{ij}) \leq D_j(1+\varepsilon) \quad \forall j$

(où $\varepsilon > 0$ est la norme de précision désirée)

on pose $T_{ij} = Q_i^n P_j^n \exp(-\theta\, c_{ij}) \quad \forall (i,j)$.

5- $$P_j^{n+1} = \text{minimum}\left(1, \frac{D_j}{\sum_i Q_i^n \exp(-\theta\, c_{ij})}\right) \forall j$$

6- $n = n+1$

retour à 3-

Cet algorithme est une généralisation de l'algorithme de balancement donné en annexe 1.

CHAPITRE III

CALIBRATION ET GÉNÉRALISATION DES THÉORÈMES DE S. EVANS

Nous allons maintenant établir les conditions qui maximisent la vraisemblance des prévisions obtenues à partir de la résolution du programme 4; un estimateur de vraisemblance maximale du paramètre θ en particulier sera développé ainsi qu'une preuve de son unicité et de son existence. Une méthode de calcul de ce paramètre concluera cette section.

Les interactions T_{ij} prévues par la résolution du programme 4 étant donné θ sont de la forme suivante:

$$T_{ij}(\theta) = \exp(-1-\alpha_i)\ \exp(-\beta_j)\ \exp(-\theta\ c_{ij}) \quad \forall(i,j)$$

et telles que:

$$\sum_j T_{ij} = O_i \quad \forall i$$

$$\sum_i T_{ij} \leq D_j \quad \forall j$$

$$\beta_j > 0 \rightarrow \sum_i T_{ij} = D_j$$

$$\sum_i T_{ij} < D_j \rightarrow \beta_j = 0$$

$$\beta_j \geq 0 \quad \forall j \ .$$

Elles doivent reproduire aussi fidèlement que possible les interactions P_{ij} observées lors de la période de calibration. Ces observations sont telles que:

$$\sum_j P_{ij} = O_i \quad \forall i$$

$$\sum_i P_{ij} = \hat{D}_j \leq D_j \quad \forall j \ .$$

L'estimateur de la vraisemblance des T_{ij} est proportionnel à $\prod_{i,j} T_{ij}^{P_{ij}}$.

Les prévisions les plus vraisemblables sont la solution du programme suivant (en maximisant le logarithme naturel):

Maximiser

$$\sum_{i,j} P_{ij}(-1-\alpha_i-\beta_j-\theta\ c_{ij})$$

sujet à

$$\sum_{i,j} \exp(-1-\alpha_i-\beta_j-\theta\ c_{ij}) = T$$

$$\beta_j \geq 0 \quad \forall j \ .$$

Formons la fonction de Lagrange:

$$L = \sum_{i,j} P_{ij}(-1-\alpha_i-\beta_j-\theta\ c_{ij}) + \lambda\{ \sum_{i,j} \exp(-1-\alpha_i-\beta_j-\theta\ c_{ij}) - T\} \ .$$

La solution de ce programme est caractérisée par:

$$\nabla_\alpha L = 0$$

$$\beta\nabla_\beta L = 0$$

$$\nabla_\theta L = 0$$

$$\nabla_\lambda L = 0$$

$$\beta \geq 0$$

$$\nabla_\alpha L = 0 \leftrightarrow -\sum_j P_{ij} - \lambda \sum_j T_{ij} = O_i \quad \forall i \ .$$

En faisant la somme sur les i , on obtient $\lambda = -1$; donc:

$$\nabla_\alpha L = 0 \leftrightarrow \sum_j T_{ij} = \sum_j P_{ij} = O_i \quad \forall i$$

$$\beta\nabla_\beta L = 0 \leftrightarrow \beta_j \sum_i T_{ij} = \beta_j \sum_i P_{ij} \quad \forall j$$

$$\therefore \beta_j > 0 \rightarrow \sum_i T_{ij} = \sum_i P_{ij} = D_j$$

$$\nabla_\theta L = 0 \leftrightarrow \sum_{i,j} c_{ij} T_{ij} = \sum_{i,j} c_{ij} P_{ij}$$

$$\nabla_\lambda L = 0 \leftrightarrow \sum_{i,j} T_{ij} = T$$

$$\beta \geq 0 \ .$$

Les T_{ij} les plus vraisemblables ont donc les caractéristiques suivantes:

$$1- \sum_j T_{ij} = O_i \quad \forall i$$

$$2- \beta_j > 0 \rightarrow \sum_i T_{ij} = \sum_i P_{ij} = D_j$$

Par conséquent, si la contrainte associée à la destination j est contraignante dans le programme 4, la destination correspondante doit être remplie à capacité dans les observations.

$$3- \sum_{i,j} c_{ij} T_{ij} = \sum_{i,j} c_{ij} P_{ij} \quad \text{ou encore} \quad \hat{C}(\theta) = \sum_{i,j} c_{ij} P_{ij} .$$

Théorème 7. $\hat{C}(\theta)$ est une fonction strictement décroissante $\forall \theta \in \mathbb{R}$.

Preuve. Montrons d'abord que $\theta^1 = \theta^2 \leftrightarrow \underset{\sim}{T}(\theta^1) = \underset{\sim}{T}(\theta^2)$. Supposons que $\underset{\sim}{T}(\theta^1) = \underset{\sim}{T}(\theta^2)$ et $\theta^1 \neq \theta^2$. Alors, il existe Q^1 , Q^2, P^1, P^2 tels que:

$$Q_i^1 P_j^1 \exp(-\theta^1 c_{ij}) = Q_i^2 P_j^2 \exp(-\theta^2 c_{ij}) \quad \forall (i,j)$$

$$\leftrightarrow \exp(-\theta^1 c_{ij}) = \frac{Q_i^2}{Q_i^1} \frac{P_j^2}{P_j^1} \exp(-\theta^2 c_{ij}) \quad \forall (i,j)$$

$$\leftrightarrow c_{ij} = u_i + v_j \quad \forall (i,j)$$

qui a été rejeté par hypothèse. Contradiction. Soient $\theta^1 \neq \theta^2$. Nous avons:

$$\psi(\underset{\sim}{T}(\theta^1),\theta^1) < \psi(\underset{\sim}{T}(\theta^2),\theta^1)$$

$$\psi(\underset{\sim}{T}(\theta^2),\theta^2) < \psi(\underset{\sim}{T}(\theta^1),\theta^2)$$

$$\hat{F}(\theta^1) + \theta^1 \hat{C}(\theta^1) < \hat{F}(\theta^2) + \theta^1 \hat{C}(\theta^2) \qquad (24)$$

$$\hat{F}(\theta^2) + \theta^2 \hat{C}(\theta^2) < \hat{F}(\theta^1) + \theta^2 \hat{C}(\theta^1) \; . \qquad (25)$$

En éliminant les $\hat{F}$ de (24) et (25), on obtient:

$$(\theta^1 - \theta^2)\,(\hat{C}(\theta^1) - \hat{C}(\theta^2)) < 0 \; . \qquad \square$$

Le théorème 7 démontre l'unicité de l'estimateur de θ .

Théorème 8. Soient:

$$m = \text{minimum} \;\; C(\underset{\sim}{T}) \;\; \text{sujet à} \;\; \underset{\sim}{T} \in \Omega$$

$$M = \text{maximum} \;\; C(\underset{\sim}{T}) \;\; \text{sujet à} \;\; \underset{\sim}{T} \in \Omega \; .$$

Alors:

$$\lim_{\theta \to \infty} \hat{C}(\theta) = m$$

$$\lim_{\theta \to -\infty} \hat{C}(\theta) = M \; .$$

Preuve. Nous allons montrer la première partie du théorème et la seconde suivra par symétrie. Soit $\overset{\circ}{\underset{\sim}{T}}$ la solution du programme de minimisation du coût global des interactions. $\overset{\circ}{\underset{\sim}{T}} \in \Omega$.

Il existe t^* et T^* tels que:

$$T^* \geq F(\underset{\sim}{T}) \geq t^* \quad \forall\, \underset{\sim}{T} \in \Omega .$$

$\hat{C}(\theta)$ est strictement décroissante et bornée inférieurement par m ; elle doit donc converger vers $m' \geq m$. Nous allons montrer que $m' = m$.

Supposons $m' > m$. Choisissons $\theta > \dfrac{T^* - t^*}{m' - m} > 0$.

Nous avons:

$$\hat{C}(\theta) \geq m' > m = C(\overset{\circ}{\underset{\sim}{T}})$$

$$t^* \leq \hat{F}(\theta) \ , \ F(\overset{\circ}{\underset{\sim}{T}}) \leq T^* .$$

D'une part;

$$\hat{F}(\theta) + \theta\, \hat{C}(\theta) < F(\overset{\circ}{\underset{\sim}{T}}) + \theta\, C(\overset{\circ}{\underset{\sim}{T}}) .$$

D'autre part,

$$F(\overset{\circ}{\underset{\sim}{T}}) + \theta\, C(\overset{\circ}{\underset{\sim}{T}}) - \hat{F}(\theta) - \theta\, \hat{C}(\theta)$$

$$= \{F(\overset{\circ}{\underset{\sim}{T}}) - \hat{F}(\theta)\} + \theta\{C(\overset{\circ}{\underset{\sim}{T}}) - \hat{C}(\theta)\} \leq (T^* - t^*) + \theta(m - m') .$$

Cette dernière quantité est inférieure à 0 par le choix de θ . Il y a donc contradiction. □

Le théorème 8 assure l'existence de l'estimateur de θ . Le calcul de cet estimateur peut se faire de plusieurs façons, dont la méthode d'approximation linéaire de Hyman (voir annexe 2).

CHAPITRE IV

EXTENSION À UNE SÉRIE DE CONTRAINTES D'INTERVALLE

Nous allons ici résoudre et calibrer le programme 5:

Programme 5.

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_{j} T_{ij} = O_i \quad \forall i \tag{26}$$

$$\sum_{i} T_{ij} \leq \overline{D}_j \quad \forall j \tag{27}$$

$$-\sum_{i} T_{ij} \leq -\underline{D}_j \quad \forall j \tag{28}$$

$$T_{ij} \geq 0 \quad \forall (i,j) \; . \tag{29}$$

Soient:

α le vecteur dual associé à l'équation (26)

β le vecteur dual associé à l'équation (27)

γ le vecteur dual associé à l'équation (28) .

La fonction de Lagrange est:

$$L(\underset{\sim}{T},\alpha,\beta,\gamma) = \sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij} + \sum_i \alpha_i (\sum_j T_{ij} - O_i)$$

$$+ \sum_j \beta_j (\sum_i T_{ij} - \overline{D}_j) - \sum_j \gamma_j (\sum_i T_{ij} - \underline{D}_j)$$

et les conditions de Kuhn-Tucker sont:

$$\overline{T}_{ij} = \exp(-1-\overline{\alpha}_i) \exp(-\overline{\beta}_j) \exp(\overline{\gamma}_j) \exp(-\theta\, c_{ij}) \quad \forall (i,j) \tag{30}$$

$$\sum_j \exp(-1-\overline{\alpha}_i) \exp(-\overline{\beta}_j) \exp(\overline{\gamma}_j) \exp(-\theta\, c_{ij}) = O_i \quad \forall i \tag{31}$$

$$\sum_i \exp(-1-\overline{\alpha}_i) \exp(-\overline{\beta}_j) \exp(\overline{\gamma}_j) \exp(-\theta\, c_{ij}) \le \overline{D}_j \quad \forall j \tag{32}$$

$$\sum_i \exp(-1-\overline{\alpha}_i) \exp(-\overline{\beta}_j) \exp(\overline{\gamma}_j) \exp(-\theta\, c_{ij}) \ge \underline{D}_j \quad \forall j \tag{33}$$

$$\overline{\beta}_j \sum_i \exp(-1-\overline{\alpha}_i) \exp(-\overline{\beta}_j) \exp(\overline{\gamma}_j) \exp(-\theta\, c_{ij}) = \overline{\beta}_j \overline{D}_j \quad \forall j \tag{34}$$

$$\overline{\gamma}_j \sum_i \exp(-1-\overline{\alpha}_i) \exp(-\overline{\beta}_j) \exp(\overline{\gamma}_j) \exp(-\theta\, c_{ij}) = \overline{\gamma}_j \underline{D}_j \quad \forall j \tag{35}$$

$$\overline{\beta}_j \ge 0 \quad \forall j \tag{36}$$

$$\overline{\gamma}_j \ge 0 \quad \forall j \; . \tag{37}$$

Le fait que les équations (27) et (28) ne soient pas indépendantes entraîne les relations suivantes:

$$\exp(\overline{\gamma}_j - \overline{\beta}_j) > 1 \iff \exp(\overline{\gamma}_j) > 1 \quad \text{et} \quad \exp(-\overline{\beta}_j) = 1 \iff \sum_i T_{ij} = \underline{D}_j \tag{38}$$

$$\exp(\overline{\gamma}_j - \overline{\beta}_j) < 1 \iff \exp(\overline{\gamma}_j) = 1 \quad \text{et} \quad \exp(-\overline{\beta}_j) < 1 \iff \sum_i T_{ij} = \overline{D}_j \tag{39}$$

Le programme dual du programme 5 est:

Programme 5^D.

Maximiser

$$- \sum_i \alpha_i O_i - \sum_j \beta_j \overline{D}_j + \sum_j \gamma_j \underline{D}_j$$

$$- \sum_{i,j} \exp(-1-\alpha_i) \exp(-\beta_j) \exp(\gamma_j) \exp(-\theta c_{ij})$$

sujet à (36) et (37).

Nous allons maintenant résoudre le programme 5^D en utilisant une méthode analogue à celle utilisée dans la section II. Soient:

$$\mathcal{T} = \mathcal{T}(\alpha,\beta,\gamma) \leftrightarrow T_{ij} = \exp(-1-\alpha_i) \exp(-\beta_j) \exp(\gamma_j) \exp(-\theta c_{ij}) \quad \forall (i,j) \quad (40)$$

$$L(\mathcal{T}(\alpha,\beta,\gamma),\alpha,\beta,\gamma) = - \sum_i \alpha_i O_i - \sum_j \beta_j \overline{D}_j + \sum_j \gamma_j \underline{D}_j$$

$$- \sum_{i,j} \exp(-1-\alpha_i) \exp(-\beta_j) \exp(\gamma_j) \exp(-\theta c_{ij})$$

$$Z = \{(\alpha,\beta,\gamma)/\alpha \in \mathbb{R}^I , \beta \in \mathbb{R}^J , \gamma \in \mathbb{R}^J , \beta \geqq 0 , \gamma \geqq 0\} .$$

Nous devons maximiser L sur l'ensemble Z. Nous allons maximiser L par rapport aux vecteurs α, β et γ alternativement en maintenant les autres vecteurs fixés tout en demeurant dans l'ensemble Z.

Soit $\beta^1 = \gamma^1 = 0$. Le maximum de L par rapport à α tel que $(\alpha,\beta^1,\gamma^1) \in Z$ est donné par:

$$1-\quad \exp(-1-\alpha_i^1) = \frac{O_i}{\sum_j \exp(-\beta_j^1)\ \exp(\gamma_j^1)\ \exp(-\theta\ c_{ij})} \quad \forall i\ .$$

Le maximum de L par rapport à β tel que $(\alpha^1,\beta,\gamma^1) \in Z$ est donné par l'expression:

$$2-\quad \exp(-\beta_j^2) = \text{minimum}\left(1,\ \frac{\overline{D}_j}{\exp(\gamma_j^1)\ \sum_i \exp(-1-\alpha_i^1)\ \exp(-\theta\ c_{ij})}\right) \quad \forall j\ .$$

Le maximum de L par rapport à γ tel que $(\alpha^1,\beta^2,\gamma) \in Z$ est donné par l'expression:

$$3-\quad \exp(\gamma_j^2) = \text{maximum}\left(1,\ \frac{\underline{D}_j}{\exp(-\beta_j^2)\ \sum_i \exp(-1-\alpha_i^1)\ \exp(-\theta\ c_{ij})}\right) \quad \forall j$$

etc...

Nous obtenons ainsi la suite:

$$\{(\alpha^n,\beta^n,\gamma^n),(\alpha^n,\beta^{n+1},\gamma^n),(\alpha^n,\beta^{n+1},\gamma^{n+1})\} \in Z \quad \forall\ n \in \mathbb{N} - \{0\}$$

telle que:

$$L(\underset{\sim}{T}(\alpha^1,\beta^1,\gamma^1),\alpha^1,\beta^1,\gamma^1) \leq L(\underset{\sim}{T}(\alpha^1,\beta^2,\gamma^1),\alpha^1,\beta^2,\gamma^1) \leq \ldots$$

L étant bornée supérieurement sur l'ensemble Z (par le théorème de dualité forte, car il existe une (et une seule) solution au programme 5 (les équations (26) à (29) définissant un ensemble convexe, fermé et borné)), cette suite converge. Après la maximisation par rapport à α, l'équation (31) est satisfaite; après la maximisation par rapport à

β , les équations (32), (34) et (36) sont satisfaites; après la maximisation par rapport à γ , les équations (33), (35) et (37) sont satisfaites. Le point vers lequel cette suite converge satisfait donc simultanément toutes les conditions de Kuhn-Tucker; il s'agit de $(\bar{\alpha},\bar{\beta},\bar{\gamma})$ et $\bar{T} = T(\bar{\alpha},\bar{\beta},\bar{\gamma})$.

En posant $Q_i = \exp(-1-\alpha_i)$, $P_j = \exp(-\beta_j)$, $R_j = \exp(\gamma_j)$ et en établissant un critère d'arrêt approprié, nous obtenons l'algorithme de résolution suivant:

1- $n = 1$

2- $P_j^n = R_j^n = 1 \quad \forall j$

3- $$Q_i^n = \frac{O_i}{\sum_j P_j^n R_j^n \exp(-\theta c_{ij})} \quad \forall i$$

4- Si $\underline{D}_j(1 - \varepsilon) < \sum_i Q_i^n P_j^n R_j^n \exp(-\theta c_{ij}) < \bar{D}_j(1 + \varepsilon) \quad \forall j$

on pose $T_{ij} = Q_i^n P_j^n R_j^n \exp(-\theta c_{ij}) \quad \forall (i,j)$

5- $$P_j^{n+1} = \text{minimum}\left(1, \frac{\bar{D}_j}{R_j^n \sum_i Q_i^n \exp(-\theta c_{ij})}\right) \quad \forall j$$

6- $$R_j^{n+1} = \text{maximum}\left(1, \frac{\underline{D}_j}{P_j^{n+1} \sum_i Q_i^n \exp(-\theta c_{ij})}\right) \quad \forall j$$

7- $n = n+1$

retour à 3-

Cet algorithme est également une généralisation de l'algorithme de balancement donné en annexe 1.

Il existe par ailleurs un autre algorithme, également apparenté à l'algorithme de balancement, pour résoudre le programme 5^D; cet algorithme a été introduit par Censor et Lent [2] et est un cas particulier d'une méthode proposée dans un article de Bregman [1].

Passons maintenant à la calibration du programme 5: il s'agit d'une question assez simple car le chapitre III est presque intégralement valable pour le programme 5 aussi. Définissons F, $\hat{F}$, C, $\hat{C}$, ψ comme dans le chapitre II. Soit P_{ij} l'interaction observée entre les zones i et j pendant la période de calibration. La solution du programme 5 la plus vraisemblable a les caractéristiques suivantes:

$$1- \quad \sum_j T_{ij} = \sum_j P_{ij} = O_i \quad \forall i$$

$$2- \quad \beta_j > 0 \rightarrow \sum_i T_{ij} = \sum_i P_{ij} = \overline{D}_j$$

$$\gamma_j > 0 \rightarrow \sum_i T_{ij} = \sum_i P_{ij} = \underline{D}_j \ .$$

Cela signifie que: si la contrainte d'intervalle est contraignante dans sa borne supérieure dans la solution du programme 5, l'observation correspondante doit être que la borne supérieure de la capacité de la destination correspondante est atteinte; si la contrainte d'intervalle est contraignante dans sa borne inférieure dans la solution du programme 5, l'observation correspondante doit être que la borne inférieure

de la capacité de la destination correspondante est atteinte.

3- $\sum_{i,j} c_{ij} T_{ij} = \sum_{i,j} c_{ij} P_{ij}$ ou encore

$$\hat{C}(\theta) = \sum_{i,j} c_{ij} P_{ij} .$$

Si on nomme Ω l'espace réalisable induit par les équations (26) à (29), les théorèmes 7 et 8 sont intégralement valables pour les résultats de cette section. Par conséquent, l'estimateur du paramètre θ existe et est unique, et peut être calculé au moyen de la méthode d'approximation linéaire de Hyman (voir annexe 2).

CHAPITRE V

RÉSULTATS NUMÉRIQUES

1. Application du programme 4.

Les méthodes proposées dans les chapitres II et III en vue de la résolution et de la calibration du programme 4 furent testées sur un modèle publié par Florian et Los [8], et dont les données provenaient d'une étude de Boyce et Desfor.

Ce modèle se situe dans un contexte de "park'n ride" et vise à prévoir l'impact de certaines politiques sur le choix des gares par les usagers de la Lindenwold High Speed Line dans la région de Philadelphie. Ce réseau comprend 7 gares et une destination finale; chaque gare est munie d'un parc de stationnement. Les variables de ce modèle ont le sens suivant:

T_{ij} = Nombre de personnes voyageant de l'origine i à la gare j

O_i = Nombre de personnes qui partent de l'origine i

D_j = Nombre de personnes que peut accueillir la gare j (tient compte de la capacité du parc de stationnement et de l'occupation moyenne des automobiles stationnées)

$T = \sum_i O_i < \sum_j D_j$ est le nombre total d'usager du réseau

c_{ij} = coût généralisé impliqué par le choix de la gare j pour un usager de la zone i (deux fonctions de coût font l'objet de tests dans l'article de Florian et Los; seule la première de ces fonctions a été retenue ici).

Les données proviennent de deux sondages faits à partir des plaques d'immatriculation des véhicules stationnés en avril et novembre 1970. La capacité totale des gares passa de 6011 à 8263 entre les deux sondages, tandis que le nombre d'origines passa de 629 à 687.

Avant de passer à la description des résultats, il convient d'énoncer ici les méthodes employées par Florian et Los pour résoudre et calibrer ce modèle, afin de pouvoir faire des comparaisons.

L'estimateur de θ qu'ils ont calculé n'était pas celui du chapitre III, mais fut choisi de la façon suivante:

Trouver θ^* tel que la solution du programme suivant:

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_{j} T_{ij} = O_i \quad \forall i$$

$$\sum_{i} T_{ij} = \hat{D}_j \quad \forall j \quad (\text{où } \hat{D}_j \text{ est le nombre observé d'usagers de la gare } j)$$

ait un coût total de déplacements égal au coût total observé.

Ils ont ensuite substitué cette valeur de θ^* dans le programme 4 et résolu ce programme au moyen de l'algorithme de Frank-Wolfe (annexe 3). Lors de la calibration, la recherche de θ^* fut faite au moyen de la méthode du nombre d'or, tandis que le programme décrit ci-dessus était lui aussi résolu au moyen de la méthode de Frank-Wolfe.

Cette méthode de calibration et de résolution fut appliquée en utilisant les données des deux sondages tour à tour, et les résultats sont donnés dans le TABLEAU 1. Le calcul de θ_1^* nécessita 5 essais et 158 secondes CP sur l'ordinateur CDC Cyber 173 de l'université de Montréal, tandis que le calcul de θ_2^* nécessita 5 essais et 181.7 secondes. La résolution du programme 4 étant donné une valeur de θ^* varia entre 57.4 et 65.6 secondes.

Avril 1970	$\theta_1^* = 6.0$
Novembre 1970	$\theta_2^* = 5.7$

TABLEAU 1: Résultats de la calibration effectuée par Florian et Los.

Au cours des tests effectués dans le cadre de ce travail, (sur l'ordinateur CDC Cyber 173 de l'université de Montréal également), les deux méthodes de calibration, soient celle de Florian et Los et celle du chapitre III, furent suivies. Un sommaire des résultats obtenus est donné

dans le TABLEAU 2.

$\hat{\theta}_1$ et $\hat{\theta}_2$ furent calibrés selon la méthode de Florian et Los. La méthode de calcul utilisée fut l'algorithme de balancement couplée avec la méthode d'approximation linéaire de Hyman. Dans les deux cas, la valeur de ε fut fixée à .001 . Le calcul de $\hat{\theta}_1$ nécessita 4 essais et 29 secondes d'exécution, tandis que le calcul de $\hat{\theta}_2$ nécessita 5 essais et 37 secondes d'exécution. On aurait normalement dû avoir $\hat{\theta}_1 = \theta_1^*$ et $\hat{\theta}_2 = \theta_2^*$; étant donné la fiabilité de l'algorithme de balancement, $\hat{\theta}_1$ et $\hat{\theta}_2$ semblent plus sûrs que θ_1^* et θ_2^* .

$\bar{\theta}_1$ et $\bar{\theta}_2$ furent ensuite calculés selon les méthodes proposées dans les chapitres II et III; i.e. au moyen de l'algorithme de balancement généralisé couplé avec la méthode d'approximation linéaire de Hyman. Le critère d'arrêt énoncé à l'étape 4- de l'algorithme du chapitre II fut remplacé par:

4- Si $P_j^n \sum_i Q_i^n \exp(-\theta\, c_{ij}) < D_j + \varepsilon \quad \forall j$ (avec $\varepsilon = 0.001$) Stop .

Un second critère d'arrêt, après la satisfaction de l'étape 4-, a également été introduit: ce critère est basé sur le théorème de dualité forte qui dit qu'à l'optimum $\psi(\bar{\mathbf{T}}) = L(\bar{\mathbf{T}},\bar{\alpha},\bar{\beta})$. Après la satisfaction de l'étape 4-, nous avons, d'après (22) et (23):

$$|\psi(\mathbf{T}) - L(\mathbf{T}(\alpha,\beta),\alpha,\beta)| \leq |\sum_j \varepsilon \ln P_j| \ .$$

Ce second critère d'arrêt est donc:

$$\left|\sum_j \varepsilon \ln P_j\right| < \delta \quad \text{où} \quad \delta > 0 .$$

Ce second critère s'est cependant révélé sans intérêt pratique: une valeur trop petite de δ entraîne une augmentation considérable du temps de résolution sans augmenter notablement la qualité de la solution.

Le calcul de $\bar{\theta}_1$ nécessita 4 essais et 16 secondes d'exécution tandis que le calcul de $\bar{\theta}_2$ nécessita 5 essais et 22 secondes d'exécution.

	AVRIL 1970	NOVEMBRE 1970
$\sum_i T_{ij} = \hat{D}_j$	$\hat{\theta}_1 = 3.389$	$\hat{\theta}_2 = 3.285$
$\sum_i T_{ij} \leq D_j$	$\bar{\theta}_1 = 3.361$	$\bar{\theta}_2 = 3.265$

TABLEAU 2: Résultats obtenus selon les deux méthodes de calibration

On remarque l'extrême similitude de tous les θ calibrés aussi bien selon le moment de référence que selon la méthode de calibration. Cela ne doit pas nous étonner: d'une part, θ est plutôt constant à court terme, et d'autre part on remarque que les deux méthodes de calibration seraient équivalentes si le modèle reproduisait parfaitement la réalité; or ce modèle donne d'excellents résultats.

Les TABLEAUX 3 et 4 donnent les résultats de prédictions de l'utilisation des parcs de stationnement selon les divers θ calibrés; toutes ces prédictions ont été exécutées en 7 secondes CP ou moins sur le CDC Cyber 173.

Une dernière remarque au sujet des résultats: les mauvais résultats des prédictions quant à la gare de Broadway sont dûs à une sous-estimation systématique des coûts pour cette station.

	Lindenwold	Ashland	Haddonfield	Westmont	Collingswood	Perry Av.	Brodway	
Capacité	1642	830	964	806	644	1000	125	
Utilisation observée	1224	765	964	733	644	866	55	
Utilisation prévue	1236	716	902	787	644	840	125	$\bar{\theta}_1$
% d'erreur	1.0	6.4	6.4	7.4	0.0	3.0	127.3	
Utilisation prévue	1227	721	903	792	644	840	125	$\bar{\theta}_2$
% d'erreur	0.2	5.8	6.3	8.0	0.0	3.0	127.3	
Utilisation prévue	1239	715	902	786	644	840	125	$\hat{\theta}_1$
% d'erreur	1.2	6.5	6.4	7.2	0.0	3.0	127.3	
Utilisation prévue	1229	720	903	791	644	840	125	$\hat{\theta}_2$
% d'erreur	0.4	5.9	6.3	7.9	0.0	3.0	127.3	

TABLEAU 3: Comparaison entre les utilisations de parcs de stationnement observées lors du premier sondage et prévues avec les paramètres $\bar{\theta}_1$, $\bar{\theta}_2$, $\hat{\theta}_1$, $\hat{\theta}_2$.

	Lindenwold	Ashland	Haddonfield	Westmont	Collingswood	Perry Av.	Brodway	
Capacité	2202	1136	1133	1108	736	1823	125	
Utilisation observée	1375	866	1079	878	736	1049	74	
Utilisation prévue	1405	809	1062	928	736	992	125	$\bar{\theta}_1$
% d'erreur	2.2	6.6	1.6	5.7	0.0	5.5	68.9	
Utilisation prévue	1394	815	1062	933	736	991	125	$\bar{\theta}_2$
% d'erreur	1.4	5.9	1.6	6.3	0.0	5.5	68.9	
Utilisation prévue	1408	807	1062	927	736	992	125	$\hat{\theta}_1$
% d'erreur	2.4	6.8	1.6	5.6	0.0	5.4	68.9	
Utilisation prévue	1397	813	1062	932	736	992	125	$\hat{\theta}_2$
% d'erreur	1.6	6.1	1.6	6.2	0.0	5.4	68.9	

TABLEAU 4: Comparaison entre les utilisation de parcs de stationnement observées lors du second sondage et prévues avec les paramètres $\bar{\theta}_1$, $\bar{\theta}_2$, $\hat{\theta}_1$, $\hat{\theta}_2$.

2. *Application du programme 5.*

L'exemple utilisé pour tester la solution et la calibration du programme 5 est tiré d'un article de Jornsten [14] sur le modèle de distribution avec contraintes d'intervalles. Il résout d'abord:

Minimiser

$$\sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_{j} T_{ij} = O_i \quad i = 1,2,3$$

$$\sum_{i} T_{ij} = D_j \quad j = 1,2,3$$

$$-\sum_{i,j} \frac{T_{ij}}{T} \ln \frac{T_{ij}}{T} = 1.82$$

$$T_{ij} \geq 0 \quad i,j = 1,2,3$$

où:

$$[c_{ij}] = \begin{pmatrix} 4 & 7 & 11 \\ 15 & 2 & 9 \\ 8 & 10 & 3 \end{pmatrix} \qquad \begin{matrix} O_1 = 30 & D_1 = 15 \\ O_2 = 30 & D_2 = 40 \\ O_3 = 30 & D_3 = 35 \end{matrix}$$

dont la solution est donnée par la matrice:

$$\begin{pmatrix} 12.2 & 12.3 & 5.5 \\ 0.4 & 24.6 & 5.0 \\ 2.4 & 3.1 & 24.5 \end{pmatrix}$$ avec coût moyen 4.655 .

Il transforme ensuite ce modèle en un modèle à contraintes d'intervalle et démontre qu'une meilleure valeur de la fonction-objectif est atteinte à partir de ce modèle modifié, en vertu d'un ensemble réalisable plus vaste.

Le modèle décrit plus haut a été repris et modifié afin de tester les méthodes énoncées dans le chapitre 5. La démarche utilisée fut la maximisation de l'entropie sous contrainte de coût plutôt que la minimisation du coût sous contrainte d'entropie. Le modèle suivant fut d'abord résolu:

Minimiser

$$\sum_{i,j} T_{ij} \ln T_{ij} + \theta \sum_{i,j} c_{ij} T_{ij}$$

sujet à:

$$\sum_{j} T_{ij} = O_i \quad i = 1,2,3$$

$$\sum_{i} T_{ij} = D_j \quad j = 1,2,3$$

$$T_{ij} \geq 0 \quad i,j = 1,2,3 \ .$$

L'algorithme de balancement fut utilisé pour la résolution et la solution optimale fut la même que celle obtenue par Jornsten, comme prévu. Nous avons donc obtenu:

$\theta = .261$

Coût moyen = 4.655

Entropie = 1.82

Temps de résolution = .181 secondes

La contrainte $\sum_i T_{ij} = D_j$ fut ensuite remplacée par: $\underline{D}_j \leq \sum_i T_{ij} \leq \overline{D}_j$. Les valeurs suivantes des vecteurs $\underline{D}$ et $\overline{D}$ furent testées:

a)
$$\underline{D} = \begin{pmatrix} 11 \\ 36 \\ 31 \end{pmatrix} \qquad \overline{D} = \begin{pmatrix} 19 \\ 44 \\ 39 \end{pmatrix}$$

Solution:

$$\begin{pmatrix} 14.5 & 11.2 & 4.3 \\ 0.6 & 24.9 & 4.5 \\ 3.9 & 3.8 & 22.3 \end{pmatrix}$$

$\theta = .2482$

Coût moyen = 4.655

Entropie = 1.85

Temps de résolution = .204 secondes

b)
$$\underline{D} = \begin{pmatrix} 12 \\ 37 \\ 32 \end{pmatrix} \qquad \overline{D} = \begin{pmatrix} 18 \\ 43 \\ 38 \end{pmatrix}$$

Solution:

$$\begin{pmatrix} 14.0 & 11.5 & 4.5 \\ 0.6 & 24.8 & 4.6 \\ 3.5 & 3.7 & 22.8 \end{pmatrix}$$

$\theta = .2504$

Coût moyen = 4.655

Entropie = 1.85

Temps de résolution = .187 secondes

c) $\underline{D} = \begin{pmatrix} 13 \\ 38 \\ 33 \end{pmatrix}$ $\overline{D} = \begin{pmatrix} 17 \\ 42 \\ 37 \end{pmatrix}$

Solution:

$$\begin{pmatrix} 13.4 & 11.7 & 4.8 \\ 0.5 & 24.7 & 4.8 \\ 3.1 & 3.5 & 23.4 \end{pmatrix}$$

$\theta = .2532$

Coût moyen = 4.655

Entropie = 1.84

Temps de résolution = .196 secondes

d)

$$\underline{D} = \begin{pmatrix} 14 \\ 39 \\ 34 \end{pmatrix} \qquad \overline{D} = \begin{pmatrix} 16 \\ 41 \\ 36 \end{pmatrix}$$

Solution:

$$\begin{pmatrix} 12.8 & 12.0 & 5.2 \\ 0.4 & 24.7 & 4.9 \\ 2.7 & 3.3 & 24.0 \end{pmatrix}$$

$\theta = .2568$

Coût moyen = 4.655

Entropie = 1.83

Temps de résolution = .208 secondes

e)

$$\underline{D} = \begin{pmatrix} 14.5 \\ 39.5 \\ 34.5 \end{pmatrix} \qquad \overline{D} = \begin{pmatrix} 15.5 \\ 40.5 \\ 35.5 \end{pmatrix}$$

Solution:

$$\begin{pmatrix} 12.5 & 12.1 & 5.3 \\ 0.4 & 24.6 & 5.0 \\ 2.6 & 3.2 & 24.2 \end{pmatrix}$$

$\theta = .2588$

Coût moyen = 4.655

Entropie = 1.82

Temps de résolution = .221 secondes

f)
$$\underline{D} = \begin{pmatrix} 14.9 \\ 39.9 \\ 34.9 \end{pmatrix} \qquad \overline{D} = \begin{pmatrix} 15.1 \\ 40.1 \\ 35.1 \end{pmatrix}$$

Solution:

$$\begin{pmatrix} 12.3 & 12.2 & 5.5 \\ 0.4 & 24.6 & 5.0 \\ 2.4 & 3.1 & 24.5 \end{pmatrix}$$

$\theta = .2606$

Coût moyen = 4.655

Entropie = 1.82

Temps de résolution = .389 secondes

Ces résultats confirment plusieurs de nos intuitions. A savoir:

1- La contrainte de coût moyen est plus contraignante lorsque l'ensemble réalisable s'amenuise (θ croît).

2- L'entropie diminue lorsque l'ensemble réalisable diminue.

3- Lorsque la longueur de l'intervalle tend vers zéro, la solution du programme tend vers la solution du programme à contraintes d'égalité.

CONCLUSION

Le présent travail démontre qu'on peut généraliser l'algorithme de balancement pour résoudre le modèle de distribution avec certaines contraintes d'inégalité ou certaines contraintes d'intervalle. Seules les relations primales-duales sont utilisées pour développer ces algorithmes et prouver leur convergence.

Il est également indiqué comment calibrer ces modèles selon la méthode du maximum de vraisemblance.

Les tests numériques prouvent en particulier de façon éclatante la supériorité de l'algorithme de balancement généralisé sur la méthode de Frank-Wolfe dans le cas du modèle à contraintes d'inégalité; un autre avantage réside dans le fait que l'algorithme de balancement généralisé résout simultanément le primal et le dual alors que la méthode de Frank-Wolfe ne résout que le primal.

Ces mêmes tests numériques ne démontrent pas de différence significative entre la méthode de calibration proposée ici et la méthode proposée par Florian et Los.

REMERCIEMENTS

Je désire remercier toutes les personnes qui m'ont aidé lors de ce travail, et en premier lieu mon directeur Marc Los. Son attention suivie et son souci du détail, de même que ses critiques pertinentes, ont contribué à rendre ce texte plus intelligible. Je demeure naturellement seul responsable des erreurs qui auraient pu persister jusqu'au texte final.

Je suis redevable au professeur Neil Stewart du département d'I.R.O. qui est à l'origine du sujet de ce travail et qui m'a donné de multiples références, et au professeur Kurt Jornsten, de l'université de Linköping, qui s'est intéressé à mes idées et qui m'a soumis le programme exposé dans le chapitre IV.

Je désire finalement remercier les employés du Service d'aide aux usagers du Centre de calcul de l'université de Montréal, qui m'ont conseillé lors des tests numériques, et Mademoiselle Thérèse Ouellet qui a bien voulu dactylographier ce texte.

BIBLIOGRAPHIE

[1] Bregman, L. (1967). The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. USSR Computational Mathematics and Mathematical Physics (7), 200-217.

[2] Censor, Y., Lent, A. (1978). An iterative row-generation method for interval convex programming. Technical report MIGP11, Medical Image Processing Group, Department of Computer Science, SUNY at Buffalo, Amherst, N.Y.

[3] Eriksson, J. (1978). Solution of large sparse maximum entropy problems with linear equality constraints. Linkoping University, Department of Mathematics. LiTH-MAT-R-77-3.

[4] Erlander, S. (1977). Entropy in linear programs - an approach to planning. Linkoping Institute of Technology, Department of Mathematics. LiTH-MAT-R-77-3.

[5] Evans, A.W. (1971). The calibration of trip distribution models with exponential or similar cost functions. Transportation Research 5, 15-38.

[6] Evans, S.P. (1973). A relationship between the gravity model for trip distribution and the transportation problem in linear programming. Transportation Research, 7, 39-61.

[7] Florian, M., Los, M. (1979). Determining intermediate origin-destination matrices for the analysis of composite mode trips. Transportation Research-B, vol. 13-B, no. 2, 91-103.

[8] Florian, M., Los, M. (1980). Impact of the supply of parking spaces on parking lot choice. A venir dans Transportation Research-B.

[9] Frank, M. et Wolfe, P. (1956). An algorithm for quadratic programming. Naval Research Logistics Quarterly, 3, 95-110.

[10] Geoffrion, A.M (1972). Duality in nonlinear programming: a simplified applications-oriented development. Dans Perspectives in Optimization, Addison-Wesley. 1972.

[11] Hyman, G.M. (1969). The calibration of trip distribution models. Environment and Planning, 1, 105-112.

[12] Jaynes, E.T. (1957). Information Theory and statistical mechanics. Phys. Rev., 106, 620-630.

[13] Jefferson, T.R., Scott, C.H. (1979). The analysis of entropy models with equality and inequality constraints. Transportation Research-B, vol. 13-B, 123-132.

[14] Jornston, K.O. (1979). A relaxed distribution model and its combined distribution and assignment conterpart. Linkoping University, Department of Mathematics, LiTH-MAT-R-79-38.

[15] Kadas, S.A., Klafsky, E. (1976). Estimation of the parameters in the gravity model for trip distribution: a new method and solution algorithm. Regional Science and Urban Economics, 6, 4.

[16] Kuhn, H.W., Tucker, A.W. (1956). Linear Inequalities and related systems. (Princeton University Press, Princeton).

[17] Mangasarian, Olvi L. Nonlinear Programming, Mc Graw-Hill, 1969.

[18] Shannon, C., Weaver, W. (1949). The mathematical theory of communication. (University of Illinois Press).

[19] Sinkhorn, R. (1967). Diagonal equivalence to matrices with prescribed row and column sums. American Mathematical Monthly, 74, 402-405.

[20] Snickars, F., Weibull, J.W. (1977). A minimum information principle: theory and practice. Regional Science and Urban Economics, vol. 7, 137-168.

[21] Theil, H. (1969). Economics and Information Theory. North-Holland, Amsterdam.

[22] Williams, I. (1976). A comparison of some calibration techniques for doubly constrained models with an exponential cost function. Transportation Research, vol. 10, 91-104.

[23] Wilson, A.G.(1970). Entropy in Urban and Regional Modelling. Pion.

[24] Wilson, A.G., Senior, M.L.(1974). Some relationship between entropy maximizing models, mathematical programming models and their duals. Journal of Regional Science, 14, 2.

ANNEXE 1

L'ALGORITHME DE BALANCEMENT

1- $n = 1$

2- $P_j^n = 1 \quad \forall j$

3- $Q_i^n = \dfrac{O_i}{\sum_j P_j^n \, B_{ij}} \quad \forall i$

4- Si $\left|D_j - \sum_i Q_i^n \, P_j^n \, B_{ij}\right| < \varepsilon \, D_j \quad \forall j$ (où $\varepsilon > 0$ est la norme de précision désirée).

On pose $T_{ij} = Q_i^n \, P_j^n \, B_{ij} \quad \forall (i,j)$

5- $P_j^{n+1} = \dfrac{D_j}{\sum_i Q_i^n \, B_{ij}}$

6- $n = n+1$

retour à 3-

ANNEXE 2

LA MÉTHODE D'APPROXIMATION LINÉAIRE DE HYMAN

Soient:

$\hat{C}(\theta^K)$ = le coût moyen des interactions de la matrice qui solutionne le programme étant donné θ^K .

C^* = le coût moyen observé et qu'on veut atteindre

1- $\theta^1 = 0.5$

2- $\theta^2 = \frac{\hat{C}(\theta^1)}{C^*}\ \theta^1$

3- Pour $n = 2,3,\ldots$ on pose:

$$\theta^{n+1} = \theta^n + \frac{(\theta^n - \theta^{n-1})(C^* - \hat{C}(\theta^n))}{(\hat{C}(\theta^n) - \hat{C}(\theta^{n-1}))}$$

On arrête lorsque $|C^* - \hat{C}(\theta^n)| < \varepsilon$ où ε est la précision désirée.

N.B. L'ajustement des θ selon l'étape 2- converge toujours, bien que lentement, alors que l'ajustement selon l'étape 3- est très rapide mais peut diverger si l'on est trop loin de la solution. Si cela se produit, il faut appliquer l'ajustement de l'étape 2- pendant quelques itérations avant de passer à l'étape 3-.

ANNEXE 3

L'ALGORITHME DE FRANK-WOLFE

Soit le problème de minimisation d'une fonction convexe sous des contraintes linéaires:

Minimiser $Z(x)$ sujet à $AX \leq b$, $X \geq 0$.

1- Etant donné une solution réalisable X^1 , poser $n = 1$

2- Trouver Y^n qui minimise
$\nabla Z(X^n)Y$, sujet à $AY \leq b$, $Y \geq 0$.

3- Soit la direction $d^n = Y^n - X^n$.
Si $|\nabla Z(X^n)d^n| \leq \varepsilon$, terminer. (Où ε est une norme de convergence appropriée et X^n est la solution optimale). Sinon,

4- Trouver la longueur du pas λ^n qui minimise
$Z(X^n + \lambda d^n)$ pour $0 \leq \lambda \leq 1$.

5- Poser
$X^{n+1} = X^n + \lambda^n d^n$
$n = n+1$
Retour à 2-